CV2001 11/12 S1

Engineering probability and statistics.

1. a)

Given:

I) $P(A) = 0.75$, $P(B) = 0.55$, $P(A \cap B) = 0.35$.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 0.75 + 0.55 - 0.35$$
$$= 0.95$$

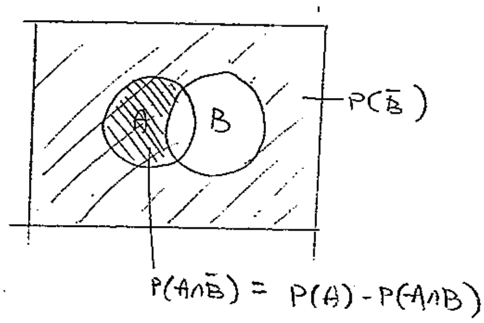
II). probability that both routes will be closed $= P(\overline{A \cup B})$

$$\therefore P(\overline{A \cup B}) = 1 - P(A \cup B)$$
$$= 1 - 0.95$$
$$= 0.05.$$

Ans: The probability is 0.05.

III)

$$P(A|\bar{B}) = \frac{P(A \cap \bar{B})}{P(\bar{B})}$$
$$= \frac{P(A) - P(A \cap B)}{1 - P(B)}$$
$$= \frac{0.75 - 0.35}{1 - 0.55}$$
$$= \frac{0.40}{0.45}$$
$$= 0.889.$$



Ans: The probability is 0.889.

IV) $P(A \cap B) = 0.35$. $P(A) = 0.75$.

$$\because P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.35}{0.55} = 0.6363$$

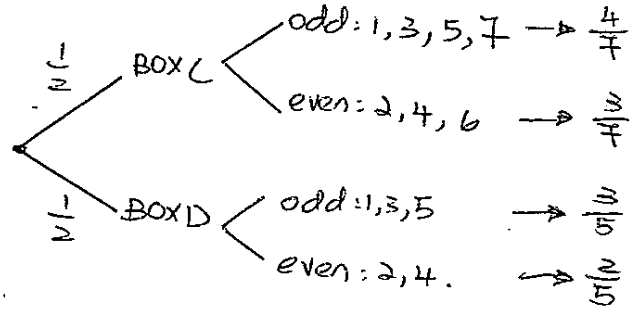
$\because P(A|B) \neq P(A)$.

$\therefore P(A \cap B) \neq P(A) + P(B)$

Ans: Events A and B are not statistically independent

1. b) let P(O) be the probability that the integer marked on the pin will be odd.

I)



* P(C) = Probability of picking box C.

$$P(O) = P(O|C).P(C) + P(O|D).P(D)$$

$$= \frac{4}{7} \times \frac{1}{2} + \frac{3}{5} \times \frac{1}{2}$$

$$= \frac{41}{70} \#$$

Ans: The probability is $\frac{41}{70}$.

II) $$P(D|O) = \frac{P(D \cap O)}{P(O)} \qquad \because P(O|D) = \frac{P(O \cap D)}{P(D)}$$

$$= \frac{P(D) \cdot P(O|D)}{P(O)}$$

$$= \frac{\frac{1}{2} \times \frac{3}{5}}{\frac{41}{70}}$$

$$= \frac{21}{41}$$

Ans: The probability is $\frac{21}{41}$.

2.

a) $P(E) \rightsquigarrow$ earthquake. $P(T) \rightsquigarrow$ typ $P(F) = \frac{1}{T}$ $\because T = 4$ yrs

$\mu = 0.05$ / yrs. $\mu = 2/10$yrs $= 0.25$.

$= 0.2$/yrs.

$\because \mu = 0.05$

$$f(x=0) = \frac{e^{-0.05} \times 0.05^0}{0!}$$

$$= 0.951229.$$

$\therefore$ The probability of at lease one earthquake accurence in a given year $= 1 - P(x=0)$

$= 0.0488.$

b) For Typhoo, $\mu = 0.4/2$yrs.

$$f(x=0) = \frac{e^{-0.4} \times 0.4^1}{1!}$$

$$= 0.2681.$$

Ans: The probability is 0.2681.

c) $P(X=3) = 5C_3 \times (P(F))^3 (P(\bar{F}))^2$.

$\because P(F) = 0.25$

$= 10 \times 0.25^3 (1-0.25)^2$

$= 10 \times 0.25^3 \times 0.75^2$

$= 0.08789.$

Ans: The probability is 0.08789.

d) $P(\text{No hazard}) = P(E=0) \times P(T=0) \times P(F=0)$

$P(E=0) \to = 0.951229.$ *

$P(T=0) \to f(x=0) = \frac{e^{-0.2} \times 0.2^0}{0!}$ $\mu = 0.2$/yrs.

$= 0.8187$ *

$P(F=0) \to P(\bar{F}) = 1 - 0.25 = 0.75$ *

$P(\text{No hazard}) = 0.951229 \times 0.8187 \times 0.75$

$= 0.584100$

$= 0.5841$ *

Ans: The probability is 0.5841.

e)

$P(E=0) = 0.95122$

$P(E=1) = \frac{e^{-0.05} \times 0.05^1}{1}$

$= 0.04756.$

$P(T=0) = 0.8187$

$P(T=1) = \frac{e^{-0.2} \times 0.2^1}{1}$

$= 0.1637$

$P(F=0) = P(\bar{F}) = 0.75$

$P(F=1) = P(F) = 0.25$

P(only one incidence of natural hazard) $= P(E=1) \cdot P(T=0) \cdot P(F=0) + P(E=0) \cdot P(T=1) \cdot P(F=0) + P(E=0) \cdot P(T=0) \cdot P(F=1)$

$= 0.04756 \times 0.8187 \times 0.75 + 0.95122 \times 0.1637 \times 0.75 + 0.95122 \times 0.8187 \times 0.25$

$= 0.6035$ #

Ans: The probability is 0.6035.

3(a)

I) Given:

$X_1 \ldots X_n \sim (\mu, \tau^2)$

$Z_1 \ldots Z_n \sim (\alpha, s^2)$

$\rho \neq 0$ when X_k, Z_i, $i = k$.

$\rho = 0$ when X_k, Z_i, $i \neq k$.

$\because \rho(X_k, Z_i) = 0$ when $k \neq i$
$= \rho$ when $k = i$.

$$Y = c \sum_{k=1}^{n} X_k + d\left[\frac{1}{n}\sum_{i=1}^{n} Z_i - \alpha\right]$$

$$E(Y) = c.E\left[\sum_{k=1}^{n} X_k\right] + \frac{d}{n}.E\left[\sum_{i=1}^{n} Z_i - \alpha\right].$$

$$\mu = c.[n.\mu] + \frac{d}{n}.[n.(\alpha - \alpha)].$$

$$\mu = c.n.\mu$$

$$(1 - cn)\mu = 0$$

$$\because 1 - cn = 0.$$

$$\therefore c = \frac{1}{n}.$$

$Var[d\alpha] = 0$

$$\therefore Var(Y) = Var\left[c\sum_{k=1}^{n} X_k + d\left(\frac{1}{n}\sum_{i=1}^{n} Z_i - \alpha\right)\right] = Var\left[c\sum_{k=1}^{n} X_k + d\left[\frac{1}{n}\sum_{i=1}^{n} Z_i\right] - d\alpha\right]$$

$$= c^2 \cdot Var.\left[\sum_{k=1}^{n} X_k\right] + \frac{d^2}{n^2} Var\left[\sum_{i=1}^{n} Z_i \ldots\right] + 2cov.\left[\sum_{k=1}^{n} X_k \sum_{i=1}^{n} Z_i\right] \times \frac{cd}{n}$$

$$= c^2 \cdot n \times Var[X_k] + \frac{d^2}{n^2} \times n.Var[Z_i] + 2\cancel{n}\rho \times \frac{c \times d}{\cancel{n}}$$

$$= c^2 \times n \times \tau^2 + \frac{d^2}{n^2} \times n \times s^2 + 2cd.\rho.$$

when Y is an unbiased estimator. $= \frac{1}{n^2} \times n \times \tau^2 + \frac{d^2}{n} \times s^2 + 2cd\rho.$

$$\cancel{0 = \frac{1}{n} \times \tau^2 + \frac{s^2 d^2}{n} + 2cd\rho.}$$

$$\cancel{0 = \frac{s^2}{n} d^2 + 2c\rho . d + \frac{\tau^2}{n}}$$

$$\cancel{d = \frac{-2c\rho \pm \sqrt{4c^2\rho^2 - 4\left(\frac{s^2}{n}\right)\left(\frac{\tau^2}{n}\right)}}{\frac{2s^2}{n}}}$$

For variance to be minimum.

$$Var(Y) = \frac{s^2}{n}d^2 + 2c\rho d + \frac{\sigma^2}{n}.$$

let $V = \frac{s^2}{n}d^2 + 2c\rho d + \frac{\sigma^2}{n}$.

- as d is the only variable.
- s, σ are constant
- ρ for 2 diff iid should be the same.

$$\frac{dv}{dd} = 0.$$

$$0 = \frac{2s^2}{n}d + 2c\rho.$$

$$\therefore d = \frac{-\cancel{2}c\rho}{\cancel{2}s^2} \times n$$

$$\boxed{d = -\frac{cn\rho}{s^2}. \qquad c = \frac{1}{n}.}$$

II). $W = \frac{1}{n}\sum_{k=1}^{n} X_k$.

$$Var[W] = \frac{1}{n^2} \cdot Var[X_k] \times n$$

$$= \frac{1}{n} \times \sigma^2$$

$$= \frac{\sigma^2}{n}.$$

$$Var[Y] = \frac{s^{\cancel{2}}}{\cancel{n}} \times \frac{-c^2n^{\cancel{2}}\rho^2}{s^{\cancel{4}2}} - 2c\rho \times \frac{cn\rho}{s^2} + \frac{\sigma^2}{n}$$

$$= \frac{\sigma^2}{n} - \frac{c^2n\rho^2}{s^2} - \frac{2c^2n\rho^2}{s^2}.$$

$$= \frac{\sigma^2}{n} - \frac{\rho^2}{ns^2} - \frac{2\rho^2}{ns^2}.$$

$\therefore$ Var[Y] consists of negative terms, hence. Var[Y] is smaller than Var[W] based on the above calculation.

Hence, ~~Var[Y]~~ The Y estimator has a lower variance.

3(b) $X_A \Rightarrow 30\%$ $\mu_A = 270\,MPa$ $\sigma_A = 5\,MPa$.

I) $X_B \Rightarrow 70\%$. $\mu_B = 275\,MPa$ $\sigma_B = 6\,MPa$.

let $Y = 0.3X_A + 0.7X_B$. be the strength of the bars.

$$E(Y) = 0.3.E(X_A) + 0.7E(X_B).$$
$$= 0.3 \times 270 + 0.7 \times 275.$$
$$= 81 + 192.5$$
$$= 273.5$$

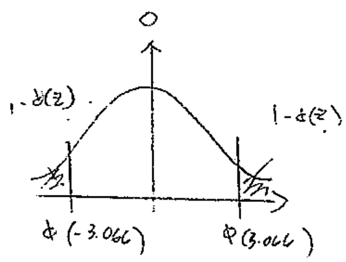
$$Var(Y) = Var[0.3X_A + 0.7X_B]$$
$$= 0.3^2 Var[X_A] + 0.7^2 var[X_B]$$
$$= 0.3^2 \times 5^2 + 0.7^2 \times 6^2$$
$$\sigma = 19.89.$$
$$cov = \frac{\sigma}{\mu} = \frac{\sqrt{19.89}}{273.5}$$
$$= 4.4598.$$

Ans: The mean of the bar is 273.5, variance is 19.89 and the coefficient of variance is 4.4598.

II) let $\bar{x}$ be the mean strength of these 100 bars and μ be the true mean of the bars.

$$\mu - \frac{0.5}{100}\mu \le \bar{X} \le \mu + \frac{0.5}{100}\mu$$
$$-\frac{0.5\mu}{100} \le \bar{X} - \mu \le \frac{0.5}{100}\mu.$$

∵ Sample true mean and variance are $(273.5, 4.4598^2)$



$$\left[\frac{\sigma}{\sqrt{n}}\right]^{-1} \times -\frac{0.5\mu}{100} \le \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} \le \frac{0.5}{100}\mu \times \left[\frac{\sigma}{\sqrt{n}}\right]^{-1}$$

$$-3.066 \le \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \le 3.066$$

follow Normal distribution.

$$\Phi(-3.066) \le z \le \Phi(3.066)$$

$1-2(1-\Phi(z))$
$1-2+2\Phi(z)$
$2\Phi(z)-1$

probability of $z = 2\Phi(3.066) - 1$
$= 2 \times 0.998893 - 1$
$= 0.997786.$

Ans: The probability that $\bar{x}$ fall within 5% of the mean strength μ is 0.9978.

4(a) $A \longrightarrow B \longrightarrow C$

normal distribution $\begin{cases} \bar{X}_A = 10 \text{ min} \\ S_A = 1.5 \text{ min} \\ n_A = 20 \end{cases}$ $\quad \bar{X}_B = 6.2 \text{ min}$, $S_B = 0.7 \text{ min}$, $n_B = 20$

let X be the total amount of time for bus number 101 from bus stop A to B, then C.

$H_0: X \leq 15 \rightarrow X - 15 \leq 0$
$H_1: X > 15 \rightarrow X - 15 > 0$.

$\therefore X = X_A + X_B$. $\quad Var[X] = Var[X_A + X_B]$

$E(X) = X_A + X_B$ $\quad = \dfrac{S_A^2}{n_A} + \dfrac{S_B^2}{n_B}$.

$= 10 + 6.2$

$= 16.2$.

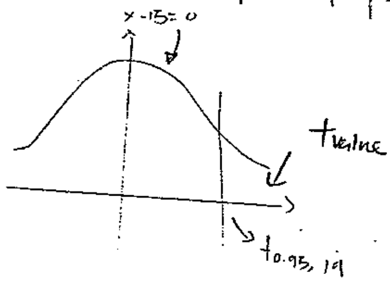
at boundary condition: $X - 15 = 0$. (we assume H_0 is true).

$$\frac{X-15}{\sqrt{\frac{S_A^2}{20} + \frac{S_B^2}{20}}} \sim t \text{ distribution.}$$

$t_{0.95,19} = 1.729$.

$$t_{value} = \frac{16.2 - 15}{\sqrt{\frac{1.5^2}{20} + \frac{0.7^2}{20}}}$$

$= 3.2420$.



$\because t_{value} > t_{0.95,19}$.

$\therefore H_0$ is rejected, we can choose to support H_1 with 95% of confidence level.

Ans: Yes, the total time taken is more than 15 minute at 95% confidence level.

4(b)

I) $\frac{d-L}{L} = \frac{F}{R}$

$Rd - RL = FL.$

$Rd = FL + RL$

$d = \frac{FL}{R} + L.$

$d = \frac{L}{R} F + L.$

$\therefore y = \hat{a}_1 x + \hat{a}_0$

$d = \frac{L}{R} F + L.$

$\therefore F \rightarrow x$

$d \rightarrow y.$

F(x) kN	d (y) mm	x^2	xy
46.4	600.294	2152.96	27853.6416
100.8	600.365	10160.64	60516.792
181.0	600.985	32761.00	108778.285
320.2	601.368	102528.04	192558.0336
$\bar{x}$ 162.1	$\bar{y}$ 600.753	Σx^2 147602.64	Σxy 389706.7522

$S_{xy} = \Sigma x_k y_k - n\bar{x}\bar{y}$

$= 389706.7522 - 4 \times 162.1 \times 600.753$

$= 178.507$

$S_{xx} = \Sigma x_k^2 - n\bar{x}^2$

$= 147602.64 - 4 \times 162.1^2$

$= 42497.$

$\therefore \hat{a}_1 = \frac{S_{xy}}{S_{xx}}$ kN mm

$=$

$= 4.200 \times 10^{-3}$

$\hat{a}_0 = \bar{Y} - \hat{a}_1 \bar{x}$

$= 600.753 - 4.200 \times 10^{-3} \times 162.1$

$= 600.072.$

$\therefore \hat{Y} = 4.200 \times 10^{-3} x + 600.072.$

$\therefore L = 600.072.$

$\frac{L}{R} = 4.200 \times 10^{-3}$

$\therefore R = 142874.3108$

$\therefore$ Parameter L is 600.072 mm and R is 142874.3108 kN.

II) $SS_T = S_{YY} = \sum_{k=1}^{4} (Y_k - \bar{Y})^2$

$= 0.793274$

$SS_R = \hat{a}_1 S_{xY} = 4.2\times10^{-3} \times 178.507$

$= 0.7497294$.

$SS_E = SS_T - SS_R$

$= 0.0435446$

$$\frac{\hat{a}_0 - a_{ref}}{\sqrt{\frac{SS_E}{n-2}\cdot\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)}} \sim \text{Normal distribution}$$

• lower bound:

$$\frac{a_{lower} - 600.072}{\sqrt{\frac{0.0435}{2}\cdot\left[\frac{1}{4} + \frac{160.1^2}{42497}\right]}} = t_{0.05,2}$$

$= -1.645$.

$$\frac{a_{lower} - 600.072}{0.137496} = -1.645$$

$a_{lower} = 599.845$.

• Upper bound:

$$\frac{a_{upp} - 600.072}{0.137496} = 1.645$$

$a_{upp} = 600.2981$.

Ans: The value between $[599.845, 600.2981]$ is 90% confidence interval.

III) We construct a test statistics.

$\therefore T_B = \dfrac{a_1 - a_0}{\sqrt{\frac{SS_E}{n-2}\times\frac{1}{S_{xx}}}}$

$H_0: a_1 = 0$

$H_1: a_1 \neq 0$.

at boundary condition.

$$|T_{boundary}| = \frac{4.2\times10^{-3} - 0}{\sqrt{\frac{0.0435446}{4-2}\cdot\frac{1}{42497}}}$$

$T_{bndry} = |5.867|$

$T_{0.95,DOF=2} = 2.920$

$\therefore H_0$ is rejected and we have 95% confident level to say that the established linear relationship has 5% of significant error.